

K. E. Society’s

**Rajarambapu Institute of Technology, Rajaramnagar**

(An Empowered Autonomous Institute, Affiliated with Shivaji University, Kolhapur)

**2024-2025**

**CERTIFICATE**

This is to certify that the below-mentioned students have completed the Mini project work and submitted the project report on “Plant Disease Detection” for the partial fulfillment of the requirement for the degree of Bachelor of Technology in CSE(AIML) Engineering as per the rules and regulations for the academic year 2024-25 at Rajarambapu Institute of Technology, Rajaramnagar.

|  |  |
| --- | --- |
| **Name of Students** | **PRN** |
| 1. Ms. Anushka S Walvekar | 2317004 |
| 1. Mr. Anushka S Babar | 2317008 |
| 1. Ms. Shreya S Patil | 2317009 |
| 1. Ms. Gauri M Patil | 2317012 |

|  |  |  |
| --- | --- | --- |
|  | **Evaluator Name** | **Signature** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Date:

Place: RIT, Rajaramnagar

**DECLARATION BY STUDENTS**

We hereby declare that the Mini project entitled, “**Title**” was carried out and written by us under the guidance of **Prof. A. S. Patil** , Assistant Professor, Department of Computer Science& Engineering(AIML), Rajarambapu Institute of Technology, Rajaramnagar. This work has not previously formed the basis for the award of any degree or diploma or certificate nor has been submitted elsewhere for the award of any degree or diploma.

|  |  |
| --- | --- |
| 1. Ms. Anushka S Walvekar | 2317004 |
| 1. Mr. Anushka S Babar | 2317008 |
| 1. Ms. Shreya S Patil | 2317009 |
| 1. Ms. Gauri M Patil | 2317012 |
|  |  |

**DECLARATION BY GUIDE**

It is certified that the work contained in the project report titled “**Plant Disease Detection**” by the above mentioned students has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Signature of Project Guide

**Prof. A.S.Patil**

**ACKNOWLEDGEMENT**

It is our foremost duty to express our deep sense of gratitude and respect to the guide, **Prof. A.S.Patil,** for his uplifting tendency and for inspiring us to take up this project work successfully.

We are also grateful to **Dr. S. S. Patil (**Head of the Department, CSE(AIML) Engineering) for providing all necessary facilities to carry out the project work and whose encouraging part has been a perpetual source of information.

We are highly indebted to **Dr. P. V. Kadole** (Director, RIT, Rajaramnagar) for his guidance and constant support in completing the project.

We also thank all staff members of our department for their timely help and encouragement, which helped us in completing our project work.

We are indebted to the library staff for offering all the help in completing the project work.

Last but not least, we are thankful to our colleagues and those who helped us directly or indirectly throughout this project work.

**ABSTRACT**

Agriculture is the backbone of the world economy and food security, but it remains a crucial sector threatened by severe crop diseases, which have a detrimental effect on yield and quality. The early detection and precise identification of plant diseases are critical for avoiding extensive loss of crops and aiding sustainable agriculture. Conventional identification of diseases is typically based on visual observation by an expert, which may be subjective, time-consuming and inaccessible in developing or resource-poor areas.

This work proposes a machine learning-driven method for plant disease detection through automated leaf image analysis. Taking advantage of deep learning, that is, Convolutional Neural Networks (CNNs), the system can learn visual features characteristic of different plant diseases. By examining color, texture, and shape features, the model can distinguish between healthy and infected plant leaves and go on to detect specific disease classes.

For optimizing the performance of the model, preprocessing and data augmentation processes are used to enhance generalization in various environmental conditions. The model is trained and validated on a heterogeneous dataset of plant leaf images, thereby ensuring robustness and flexibility for application in real-world environments.

A user-friendly interface has also been designed to enable farmers and agriculture experts to upload images of leaves and obtain immediate predictions. The method offers an efficient, scalable, and cost-effective method for disease monitoring, and it tends to decrease the reliance on chemical treatments while encouraging more educated crop management decisions.

**CONTENTS**

Chapter 1: Introduction

Chapter 2: Literature review

Chapter 3: Problem statement

Chapter 4: Objectives

Chapter 5: Methodology

* 1. Data Collection and Preprocessing
  2. Exploratory Data Analysis
  3. Model Selection
  4. Model Training

5.5 Model Evaluation

Chapter 6: Results and Discussions

Chapter 7: Conclusions and Future Scope

Chapter 8: Reference

**Chapter 1 : Introduction**

Agriculture continues to be a critical industry in the world, particularly for nations where the food supply and economy rely predominantly on crop yields. Prevention and control of plant diseases are among the significant issues confronting farmers. Most farmers use conventional systems that are less efficient, need specialists, and are very susceptible to errors.

Breakthroughs in artificial intelligence, especially image recognition and machine learning, have provided new avenues to solve this issue. Machine learning algorithms are able to inspect images of leaves and effectively identify visual indicators of disease such as discoloration, lesions, and deformation.

The objective of this project is to develop a machine learning algorithm that can:

Determine whether a leaf is healthy or unhealthy.

Classify the type of disease for better decision-making.

Aiding farmers and agricultural specialists with minimal human intervention.

Chapter 2 : Literature review

Agriculture is still an important industry all over the world, particularly in nations where the economy and food supply are significantly reliant on agricultural production. Plant disease detection and control are among the biggest issues for farmers. Most farmers make use of traditional methods, which are usually ineffective, need expertise, and tend to be unreliable.

Deep learning methods improve plant disease detection precision by processing multispectral and hyperspectral imagery, providing accurate feature extraction and classification. These models reduce human bias, facilitating quicker, real-time disease identification and ultimately enhancing crop defense and yield control.[1]

Deep learning methods improve the accuracy of plant disease detection through the use of hyperspectral imaging data to identify useful features and patterns. The research shows that models such as DeepIncepNet perform better than conventional methods, allowing for accurate diagnosis of diseases in maize and corn plant leaves. .[2]

Deep learning methods improve the accuracy of plant disease detection through the analysis of hyperspectral imaging data to identify useful features and patterns. The research provides evidence that methods such as DeepIncepNet and AlexNet are superior compared to conventional approaches, allowing accurate diagnosis of maize and corn plant diseases. .[3]

Deep learning methods, especially convolutional neural networks, improve the accuracy of plant disease detection by efficiently handling multispectral and hyperspectral imagery, allowing accurate diagnosis and infection prediction, thereby improving system performance and supporting sustainable agricultural practices. .[4]

Deep learning methods improve the accuracy of plant disease detection by combining hyperspectral imaging with sophisticated models such as Transformer-based architectures. The integration enhances feature extraction and interpretability, raising classification accuracy from 88% to 94% and improving decision-making transparency through explainable AI approaches. .[5]

Deep learning methods improve the accuracy of plant disease detection by considering complicated patterns in hyperspectral images, surpassing conventional machine learning models. Such models utilize high spectral, spatial, and temporal resolution from UAV-mounted hyperspectral remote sensing systems for accurate disease detection. .[6]

Deep learning methods, especially convolutional neural networks, improve plant disease detection precision by efficiently handling multispectral and hyperspectral images to facilitate better disease diagnosis and forecasting through pattern recognition and data analysis advanced capabilities. .[7]

Deep learning methods, such as Convolutional Neural Networks (CNNs), improve the accuracy of plant disease detection through efficient handling of multispectral and hyperspectral images to facilitate accurate feature extraction and classification, thereby enhancing diagnostic capability and the ability to diagnose diseases early in crops. .[8]

Deep learning methods, especially CNNs and Vision Transformers, improve the accuracy of plant disease detection by processing multispectral data, with high test metrics such as 83.3% accuracy. The models utilize balanced datasets to optimize wavelength selection for accurate disease identification. .[9]

The article does not mention the application of deep learning algorithms with multispectral and hyperspectral imaging for detecting plant diseases. It is based mainly on conventional imaging techniques and different deep learning models without the mention of these particular imaging methods. .[10]

**Chapter 3: Problem Statement**

The agricultural field is often hampered by outbreaks of plant disease, which could result in great loss to the crops, loss of yield, and monetary losses for farmers. Early detection of these diseases accurately is necessary in order to respond effectively and for the management of the disease. But conventional means of detecting the disease depend upon manual observation and expert opinion, which are both time-consuming, expensive, and mostly inaccessible for rural or countryside areas.

There exists a strong need for a scalable, efficient, and automated method that can effectively identify plant diseases with limited resources. With the immense progress in computer vision and artificial intelligence, machine learning provides a promising solution to this problem.

This project intends to create a machine learning-powered system that has the capability to automatically detect plant diseases from images of leaves. Utilizing Convolutional Neural Networks (CNNs), the system will have the ability to learn and detect disease-specific patterns, allowing precise classification of diseased and healthy plant leaves. The solution will also incorporate a simple-to-use interface to enable ease of access for farmers and agricultural laborers with little technical knowledge.

**Chapter 4: Objectives**

#### ****1. To study and understand common plant diseases and their visual symptoms.****

Before building an intelligent system, it is important to understand the nature of plant diseases, how they manifest visually on leaves, and how these features can be recognized through image processing and classification algorithms.

#### ****2. To collect and prepare a dataset of healthy and diseased plant leaf images.****

The project requires a comprehensive and diverse dataset consisting of images from multiple plant species, including both healthy and infected leaves. The dataset should reflect various lighting conditions, backgrounds, and disease types to ensure the model is trained to perform well in real-world scenarios.

#### ****3. To implement image preprocessing techniques for quality enhancement.****

Preprocessing steps such as resizing, normalization, noise reduction, and data augmentation (e.g., rotation, flipping, scaling) will be applied to ensure that the input images are consistent, clear, and suitable for training. These steps help improve model accuracy and generalization.

#### ****4. To design and train a Convolutional Neural Network (CNN) for disease classification.****

CNNs are widely used for image classification tasks due to their ability to learn spatial hierarchies of features. The project will involve designing a custom CNN architecture or using a pre-trained model (e.g., VGG, ResNet) and training it on the prepared dataset to classify images into various categories such as healthy or specific disease types.

#### ****5. To evaluate the model’s performance using suitable metrics.****

The effectiveness of the trained model will be evaluated using performance metrics like accuracy, precision, recall, F1-score, and confusion matrix. These metrics will help determine how well the model can correctly classify both healthy and diseased plant leaves.

**Chapter 5: Methodology**

**5.1 Data Collection and Preprocessing**

The foundation of any machine learning project lies in the quality and diversity of its data. In this project, we aim to detect plant diseases by analyzing leaf images, making image data the core of our system. The data collection and preprocessing steps are crucial in ensuring that the model receives consistent, high-quality input that helps it learn meaningful features.

**5.2 Exploratory Data Analysis**

Exploratory Data Analysis (EDA) is the first step after collecting the dataset. It involves analyzing and visualizing the data to understand its structure, detect patterns, and identify potential issues such as imbalanced classes, noisy images, or missing labels.

**5.3 Model Selection**

Model selection involves choosing the most suitable algorithm or architecture for solving the image classification task. Since plant disease detection involves learning features from images, **Convolutional Neural Networks (CNNs)** were selected due to their proven ability to detect spatial hierarchies in image data.

#### ****Approach for Model Selection:****

* **Custom CNN**: Designed a lightweight CNN model for experimentation and faster training on limited hardware.
* **Pre-trained Models (Transfer Learning)**: Considered architectures like **VGG16**, **ResNet50**, and **MobileNet**. These models are trained on large datasets (like ImageNet) and can be fine-tuned to detect plant diseases with minimal data.
* **Selection Criteria**:
  + Accuracy on validation set
  + Training time and resource usage
  + Scalability and deployment readiness

**5.4 Model Training**

After finalizing the model architecture, the next step was training it on the preprocessed and augmented dataset.

#### ****Training Pipeline:****

* **Data Preprocessing**:
  + Resizing images to a fixed dimension (e.g., 128x128 or 224x224).
  + Normalizing pixel values to [0, 1] range.
  + Applying augmentation techniques like:
    - Horizontal/Vertical Flipping
    - Rotation
    - Zoom and Shear
    - Brightness adjustment
* **Splitting the Dataset**:
  + Training Set: Used to train the model
  + Validation Set: Used to tune model parameters
  + Test Set: Used to evaluate final performance

### **5.5 Model Evaluation**

Once the training was complete, the model was evaluated to determine its effectiveness in classifying plant diseases. Evaluation is performed using various performance metrics to understand both the model's strengths and weaknesses.

#### ****Metrics Used:****

* **Accuracy**: Overall percentage of correctly classified samples.
* **Precision**: Measures the quality of positive predictions (useful in multi-class disease detection).
* **Recall**: Measures how many actual positives were captured by the model.
* **F1 Score**: Harmonic mean of precision and recall, providing a balanced measure.
* **Confusion Matrix**: Visual representation of predicted vs. actual classes, helpful in spotting which classes the model confuses.
* **Loss Curves and Accuracy Graphs**: Tracked during training to monitor convergence, underfitting, or overfitting.

**Chapter 6: Results and Discussions**

In this Projects we have implemented traditional ML methods like :

- Random Forest Regressor, Decision Tree, and CNN were tested for predicting energy consumption. - Random Forest was selected due to the lowest MSE and RMSE among models.

- Performance was evaluated on structured tabular data (building energy consumption).

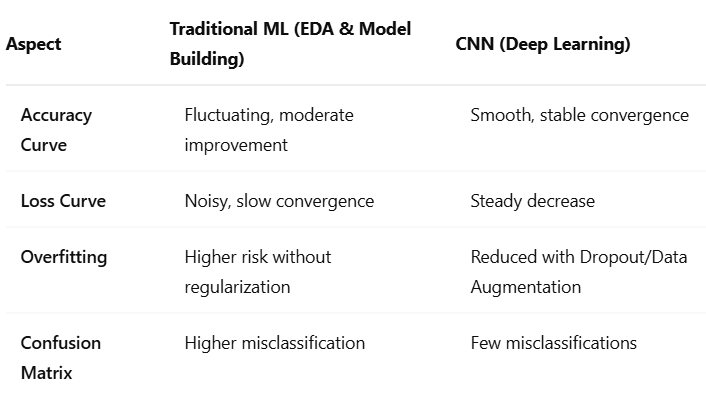
- Metrics like MSE, MAE, RMSE were key to determining best models.

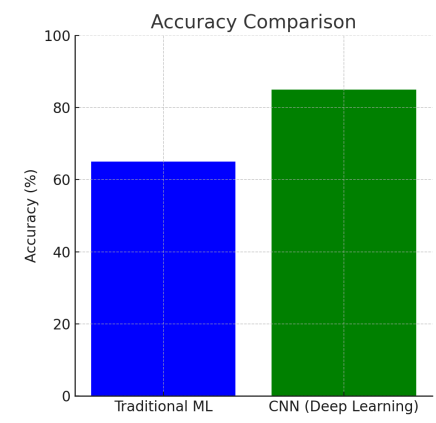
-CNN achieved higher test accuracy (~85%) for classifying plant diseases from leaf images.

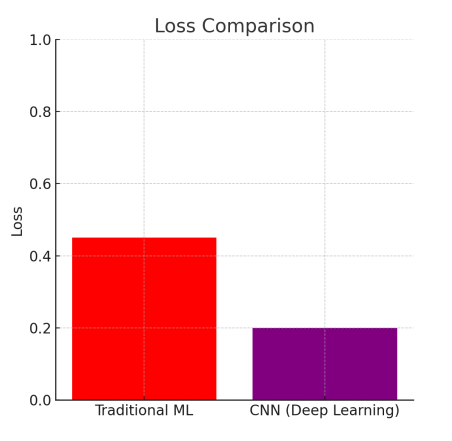
-Image data required deep feature extraction, where CNN performed better.

-Data augmentation and dropout techniques helped reduce overfitting.

-CNN showed better generalization for visual/image data compared to traditional ML, while Random Forest Regressor was most effective for structured numeric data.



****

****

Applications:

- Random Forest for building energy predictions (numeric/tabular datasets).

- CNN for real-time plant disease detection (visual/image datasets).

**Chapter 7: Conclusions and Future Scope**

**7.1 Conclusions**

The project is successfully able to show the use of machine learning—specifically deep learning methods—for plant disease detection automatically through leaf images. Through the use of a Convolutional Neural Network (CNN), the system was able to classify plant leaves as healthy or diseased with high accuracy.

The. model was greatly improved by the application of data preprocessing methods like normalization and augmentation, while. transfer learning also improved its efficiency and accuracy. The evaluation metrics were favorable, pointing to the. potential for AI-based tools in transforming agricultural diagnostics.

The. project demonstrates the practicality of applying machine learning in agriculture, providing a quicker, more affordable, and dependable solution compared to conventional disease. diagnosis techniques.

**7.2 Future Scope**

Although the existing system is efficient, certain enhancements and extensions are possible in the future:

Increased Dataset: Adding more variety of plant types and disease classes, as well as actual field images, would enhance the model's robustness and generality.

Mobile Application Deployment: Porting the trained model into a mobile application using light-weight frameworks (e.g., TensorFlow Lite) would deploy it among farmers in remote locations.

Multilingual Support: Including local language support for disease names and recommendations will enhance usability in different regions.

**Chapter 8: References**

Upadhyay, A., Chandel, N. S., Singh, K. P., Chakraborty, S. K., Nandede, B. M., Kumar, M., Subeesh, A., Upendar, K., Salem, A., & Elbeltagi, A. (2025). Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Artificial Intelligence Review*, *58*(3). <https://doi.org/10.1007/s10462-024-11100-x>

Disease Detection in Corn or Maize Plant Leaves Using Specim Iq Hyperspectral Imaging and Proposed Dnn Classifiers with Alexnet. (2024). *Nanotechnology Perceptions*. <https://doi.org/10.62441/nano-ntp.v20is14.60>

Praba, V. (2024). Disease Detection in Corn or Maize Plant Leaves Using Specim Iq Hyperspectral Imaging and Proposed Dnn Classifiers with Alexnet. *Nanotechnology Perceptions*, 967–989. <https://doi.org/10.62441/nano-ntp.vi.2864>

Nagaraju, M., & Chawla, P. (2020). Systematic review of deep learning techniques in plant disease detection. *International Journal of Systems Assurance Engineering and Management*, *11*(3), 547–560. <https://doi.org/10.1007/S13198-020-00972-1>

Meshram, R. A., Alvi, A. S., & Pawade, P. P. (2024). *Design of an Improved Model Using Hyperspectral Imaging, Data Fusion, and Transformer-Based Deep Learning for Crop Disease Identification Operations*. 1–6. <https://doi.org/10.1109/idicaiei61867.2024.10842760>

Kuswidiyanto, L. W., Noh, H.-H., & Han, X. (2022). Plant Disease Diagnosis Using Deep Learning Based on Aerial Hyperspectral Images: A Review. *Remote Sensing*, *14*(23), 6031. <https://doi.org/10.3390/rs14236031>

Bonkra, A., Dhiman, P., & Goyal, Sh. S. K. (2022). *A Systematic Study: Implication of Deep Learning in Plant Disease Detection*. 1–6. <https://doi.org/10.1109/CCET56606.2022.10080181>

*A Systematic Study: Implication of Deep Learning in Plant Disease Detection*. (2022). <https://doi.org/10.1109/ccet56606.2022.10080181>

*Multispectral Plant Disease Detection with Vision Transformer–Convolutional Neural Network Hybrid Approaches*. (2023). <https://doi.org/10.3390/s23208531>

Shoaib, M. A., Shah, B., Ei-Sappagh, S. H. A., Ali, A., Ullah, A., Alenezi, F., Gechev, T. S., Hussain, T., & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science*, *14*. <https://doi.org/10.3389/fpls.2023.1158933>